**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/81, C07K 15/00, C12N 15/62, C07K 13/00, C12N 1/21, 1/19, 5/10, C12P 21/08, A61K 39/00, G01H 33/569

A1

(11) International Publication Number: WO 94/14969

(43) International Publication Date: 7 July 1994 (07.07.94)

(21) International Application Number: PCT/GB93/02656

(22) International Filing Date: 24 December 1993 (24.12.93)

(30) Priority Data:
9227068.5 29 December 1992 (29.12.92) GB

(71) Applicant (*for all designated States except US*): BRITISH BIO-TECHNOLOGY LIMITED [GB/GB]; Watlington Road, Cowley, Oxford 0X4 5LY (GB).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): ADAMS, Sally, Elizabeth [GB/GB]; British Bio-Technology Limited, Watlington Road, Cowley, Oxford OX4 5LY (GB). BURNS, Robert, Nigel [GB/GB]; British Bio-Technology Limited, Watlington Road, Cowley, Oxford OX4 5LY (GB). RICHARDSON, Simon, Mark, Harold [GB/GB]; British Bio-Technology Limited, Watlington Road, Cowley, Oxford OX4 5LY (GB).

(74) Agent: WALLS, Alan, J.; British Bio-technology Limited, Watlington Road, Cowley, Oxford OX4 5LY (GB).

(81) Designated States: AU, CA, CZ, DE, FI, GB, HU, JP, KR, NO, NZ, RU, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: NOVEL PROTEINACEOUS PARTICLES

(57) Abstract

A non-natural particle-forming protein comprising a self-assembling particle-forming first amino acid sequence substantially homologous with a yeast retrotransposon Ty p1 protein and a second amino acid sequence, wherein the second sequence is antigenic and is incorporated within an epitope of the first amino acid sequence, which epitope, on particles formed from the first amino-acid sequence alone, is surface-exposed.

***FOR THE PURPOSES OF INFORMATION ONLY***

**Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.**

| | |
|---|---|
| AT | Austria |
| AU | Australia |
| BB | Barbados |
| BE | Belgium |
| BF | Burkina Faso |
| BG | Bulgaria |
| BJ | Benin |
| BR | Brazil |
| BY | Belarus |
| CA | Canada |
| CF | Central African Republic |
| CG | Congo |
| CH | Switzerland |
| CI | Côte d'Ivoire |
| CM | Cameroon |
| CN | China |
| CS | Czechoslovakia |
| CZ | Czech Republic |
| DE | Germany |
| DK | Denmark |
| ES | Spain |
| FI | Finland |
| FR | France |
| GA | Gabon |
| GB | United Kingdom |
| GE | Georgia |
| GN | Guinea |
| GR | Greece |
| HU | Hungary |
| IE | Ireland |
| IT | Italy |
| JP | Japan |
| KE | Kenya |
| KG | Kyrgystan |
| KP | Democratic People's Republic of Korea |
| KR | Republic of Korea |
| KZ | Kazakhstan |
| LI | Liechtenstein |
| LK | Sri Lanka |
| LU | Luxembourg |
| LV | Latvia |
| MC | Monaco |
| MD | Republic of Moldova |
| MG | Madagascar |
| ML | Mali |
| MN | Mongolia |
| MR | Mauritania |
| MW | Malawi |
| NE | Niger |
| NL | Netherlands |
| NO | Norway |
| NZ | New Zealand |
| PL | Poland |
| PT | Portugal |
| RO | Romania |
| RU | Russian Federation |
| SD | Sudan |
| SE | Sweden |
| SI | Slovenia |
| SK | Slovakia |
| SN | Senegal |
| TD | Chad |
| TG | Togo |
| TJ | Tajikistan |
| TT | Trinidad and Tobago |
| UA | Ukraine |
| US | United States of America |
| UZ | Uzbekistan |
| VN | Viet Nam |

# NOVEL PROTEINACEOUS PARTICLES

The present invention relates to biologically useful particles. In particular it relates to modified particles derived from the yeast retrotransposon Ty. Particles formed from such proteins are immunogenic and can be used in immunotherapeutic or prophylactic vaccines or as diagnostic agents.

An ideal immunogen is a polymer of multiple antigen determinants assembled into a high molecular weight, particulate complex. A substantial disadvantage of most antigens produced by recombinant DNA techniques for vaccines is that they are usually produced as simple monomeric proteins. This is not the ideal configuration of an immunising antigen as it does not readily permit the cross-linking of the components of the immune system. Such crosslinking is required for maximum stimulations of humoral and cellular immunity. For these reasons it would be advantageous to develop polyvalent, particulate carrier systems for immunising antigens.

WO-A-8803562 and WO-A-8803563 describe the use of certain fusion proteins derived from retrotransposons or RNA retroviruses for pharmaceutical, diagnostic or purification applications. Such particles are designated virus-like particles (VLPs) when derived from the yeast retrotransposon Ty. The above published PCT applications note that polyvalent particles are useful for immunisation purposes because their polyvalent nature provides that more antibodies will be raised against the particulate antigens used. The particles are formed of fusion proteins having a particle-forming sequence and, in some embodiments at least, an antigenic sequence. In the examples, the antigenic sequence is positioned C-terminal to the particle-forming sequence.

While the above approach is promising, a potential difficulty is that insertion of the antigen at the C-terminal end of the particle-forming protein may not in all cases be optimal for presentation to the immune system. Animals immunised with recombinant VLPs may elicit a higher titre response to the Ty component than to the added antigen. It would therefore be highly advantageous to construct antigen-presenting particles where the antibody response to the added antigen is augmented. Such particles might also have enhanced ability to stimulate a cell-mediated immune

response, such as a T-cell response, a Cytotoxic T-lymphocyte (CTL) reponse or a Natural Killer (NK) cell response. It would further be advantageous if, following immunisation with such particles, the antibody response to the particle-forming moiety was reduced or preferably prevented.

One way to improve the presentation of the antigenic sequence to the immune system might be to insert the antigenic sequence of interest within the particle-forming sequence. However, correct insertion of the antigenic site within the particle-forming protein is likely to be critical for particle formation. Insertions might disrupt the secondary and tertiary structure determinants of the monomer, or the quaternary interactions between monomers necessary for particle formation.

One approach to deduce suitable surface-exposed insertion sequences has been to use the understanding of the three-dimensional structure of viruses elucidated by X-ray crystallography. Such precise analysis of the structure of the polio virus has enabled particulate chimaeric proteins to be created whereby heterologous antigenic sequences are substituted for amino-acids present in the suface-exposed epitopes of this virus (Dedieu *et al.*, J. Virol. (1992) 66 3161-3167; Burke *et al*, Nature (1988) 332 81-82; Evans *et al.*, Nature (1989) 339 385-388). However, these polio virus constructions are limited by the need to produce a viable virus; even some very short sequences cannot be tolerated.

Detailed analysis as described for poliovirus is not possible for proteins which have not yet been crystallised. Where particles have a well-characterised tertiary ß-barrel structure, internal insertions of heterologous antigenic sequences into presumed surface exposed regions have been made using predictive models based on sequence alignment. For example, hybrid particles prepared from the hepatitis B core antigen and an antigen derived from a virus with an analogous secondary structure were found to maintain particle formation and enhance the immunogenicity of the inserted antigen (Schodel *et al.*, J. Virol. (1992) 66 106-114; Brown *et al.*, Vaccine 1991 9 595-601). Substitutions of heterologous peptides into presumed surface-exposed, immunodominant regions of the hepatitis B surface antigen also gave rise to particulate, chimaeric proteins with enhanced immunogenicity (von Brunn *et al.*, Vaccine 1991 9 477-601), although considerable amounts of lipid were found to be associated.

However, retrotransposons have a very poorly understood structure and it is not currently believed that they possess a ß-barrel (Burns *et al.*, J. Mol. Biol. (1990) 261 207-211). Suitable sites for insertion of antigens into these particulate proteins are therefore not known or predictable. In retroviruses (which have a very similar structure to retrotransposons) it has been shown that insertion of an antigen into the middle of the gag sequence destroys the particle-forming nature of this sequence (Luo *et al.*, Proc Natl. Acad. Sci. USA 89 10527-10531 (1992)).

The present inventors have identified the surface-exposed immunodominant epitopes within the yeast retrotransposon Ty p1. Immunogenic sites are not necessarily surface exposed; high titre antibodies are frequently elicited against core proteins during viral infections even though such proteins are not exposed on the surface of the particle (eg the influenza nucleoprotein). The inventors have also found that insertion of heterologous antigenic sequences into such epitopes does not prevent particle formation. In retrotransposons the size of insertion which can be tolerated without disrupting particle formation appears to be remarkably large; much greater than has been described for any other system, where generally substitutions have been preferred. The resulting hybrid particles exhibit reduced immunogenicity of the particle forming protein, and an enhanced immune reponse to the inserted sequence.

According to a first aspect of the invention, there is provided a non-natural particle-forming protein comprising a first self-assembling particle forming amino acid sequence substantially homologous with a yeast retrotransposon Ty p1 protein and a second amino acid sequence, wherein the second sequence is antigenic and is incorporated within an epitope of the first amino acid sequence, which epitope, on particles formed from the first amino-acid sequence alone, is surface-exposed.

Such constructions may be produced either by insertion of antigenic sequences into these surface epitopes to form true hybrid proteins, or by substitution of the native amino acids found at such sites with the amino acid sequence of interest, or by a combination of deletion, substitution and insertion.

The surface-expressed epitopes will generally be found in the N-terminal half of the first particle forming protein, the sequence of which is disclosed in Dobson *et al.*, 1984 EMBO J. 3 1115. In particular, three consensus surface-exposed regions have

been identified in the N-terminal half of the particle-forming protein p1 of the retrotransposon Ty, located at amino acids 21-42 (position A), amino acids 55-74 (position B) and amino acids 93-142 (position C) as shown in Figure 1 and summarised in Table 1. Proteins in which the second amino acid sequence is located within at least one of these regions in the first amino acid sequence are preferred. Within these regions, any suitable insertion site may be chosen for the second sequence. These sites include those between amino-acids 30-31, 67-68, 113-114 and 132-133 of the Ty protein and have been designated sites A, B, $C_1$ and $C_2$ respectively, but other sites are equally appropriate.

Particles derived from Ty may have advantages over those derived from polio or Hepatitis for use as vaccines. Pre-exposure to hepatitis or polio vaccines can compromise an effective subsequent reponse against the chimaera. The use of particles derived from Ty is therefore preferable, as there will be less likelihood of a patient having a pre-existing immunological response. Since Ty is not a pathogen, vaccination with Ty will not cause exposure to pathogenic antigens.

The expression "substantially homologous", when describing the relationship of an amino acid sequence to a natural protein, means that the amino acid sequence can be identical to the natural protein or can be an effective but truncated or otherwise modified form of the natural protein or can share at least 50%, 60%, 70%, 80%, 90%, 95% or 99%, in increasing order of preference, of the residues of the natural protein or its modified form. "Effective" means that the particle forming ability of the natural protein is retained (or at least not substantially lost). Alternatively or in addition, a nucleic acid sequence encoding the amino acid sequence may hybridise, for example under stringent conditions, to a nucleic acid sequence encoding the natural protein or its truncated form, or would do so but for the degeneracy of the genetic code. Stringent hybridisation conditions are known and are exemplified by approximately 0.9 molar salt concentration at approximately 35° to 65°C.

The antigenic sequence may correspond to a sequence derived from or associated with an aetiological agent or a tumour. The aetiological agent may be a microorganism such as a virus, bacterium, fungus or parasite. The virus may be: a retrovirus, such as HIV-1, HIV-2, HTLV-I, HTLV-II, HTLV-III, SIV, BIV, LAV,

ELAV, CIAV, murine leukaemia virus, Moloney murine leukaemia virus, and feline leukaemia virus; an orthomyxovirus, such as influenza A or B; a paramyxovirus, such as parainfluenza virus, mumps, measles, RSV and Sendai virus; a papovavirus, such as HPV; an arenavirus, such as LCMV of humans or mice; a hepadnavirus, such as Hepatitis B virus; a herpes virus, such as HSV, VZV, CMV, or EBV. The tumour-associated or derived antigen may for example be a proteinaceous human tumour antigen, such as a melanoma-associated antigen, or an epithelial-tumour associated antigen such as from breast or colon carcinoma.

The antigenic sequence may be also derived from a bacterium, such as of the genus *Neisseria*, *Campilobacter*, *Bordetella*, *Listeria*, *Mycobacteria* or *Leishmania*, or a parasite, such as from the genus *Plasmodium*, especially *P. falciparum*, or from a fungus, such as from the genus *Candida*, *Aspergillus*, *Cryptococcus*, *Histoplasma* or *Blastomyces*.

The antigenic sequence may typically vary in length from between 6 and 60 amino acids, for example 6-50, 6-40, or 6-30, although it is not possible with precision to give universally appropriate maxima and minima. The sequence should be sufficiently long to give rise to the desired immunogenic response, but not so long as to cause unacceptable distortion to the rest of the molecule.

Preferred antigenic sequences are antigenic sequences corresponding to epitopes from a retrovirus, a paramyxovirus, an arenavirus or a hepadna virus, or a from human tumour cell. Examples include known epitopes from:

1) HIV (particularly HIV-1) gp120,
2) HIV (particularly HIV-1) p24,
3) Influenza virus nucleoprotein and haemagglutinin,
4) LCMV nucleoprotein,
5) HPV L1, L2, E4, E6 and E7 proteins,
6) p97 melanoma associated antigen,
7) GA 733-2 epithelial tumour-associated antigen,
8) MUC-1 epithelial tumour-associated antigen,
9) Mycobacterium p6,
10) Malaria CSP or RESA antigens,
11) VZV gpI, gpII or gpIII

Particularly preferred antigenic sequences comprise a sequence substantially homologous with an antigenic portion of the third variable domain of a lentivirus. This region, known as the V3 loop or GPGR loop is found between amino acids 300 and 330 of the envelope glycoprotein gp120 of HIV-1 and in analogous positions of other lentiviruses. The V3 loop is defined by two flanking cysteine residues linked by a disulphide bond and, for HIV-1 at least, is the major neutralising epitope of the virus (Putney *et al* 1986 *Science* **234**, 1392; Rusche *et al* 1988 *Proc. Natl. Acad. Sci.* **85**, 3198; Palker *et al* 1988 *Proc. Natl. Acad. Sci.* **85** 1932; and Goudsmit *et al* 1988 *AIDS* **2** 157). The antigenic portion of choice may constitute the whole or half of the V3 loop. However, a conserved sequence of the V3 loop may be useful in conferring immunity against more than one isolate of a virus (such as HIV-1).

A number of isolates of HIV-1, in which the sequence of the V3 loop varies from isolate to isolate, are known. The most common isolates are HXBII, RF and MN; MAL, ELI and BH10 are also important, but the MN isolate may be the most clinically relevant. Laboratory isolate IIIB is a mixture of strains BH10 and HXBII. The invention is not limited to sequences derived from the V3 loop of any particular isolate, some of which are shown below.

| | |
|---|---|
| BH10 | SNCTRPNNNTRKSIRIQRGPGRAFVTIGKIGNMRQAHCNISG |
| HXBII | SNCTRPNNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHCNISG |
| MN | SNCTRPNYNKRKRIHIGPGRAFYTTKNIIGTIRQAHCNISG |
| MAL | SNCTRPGNNTRRGIHFGPGQALYTTGIVDIRRAYCTING |
| RF | SNCTRPNNNTRKSITKGPGRVIYATGQIIGDIRAHCNLSGS |
| ELI | STCARPYQNTRQRTPIGLGQSLYTTRSRSIIGQAHCNISG. |

Neither is the invention limited to natural V3 loop sequences. Examples of variant V3 loop sequences which can be used in the invention include:

| | |
|---|---|
| MAL(var) | SNCTRPGNNTRRGIHFGPGQALYTTGIVDEIRRAYCNISG |
| RF(var) | SNCTRPNNNTRKSITKQRGPGRVLYATGQIIGDIRKAHCNSIG |
| ELI(var) | STCARPYQNTRQRTPIGLGQSLYTTRGRTKIIGQAHCNISG. |

A comparison of the sequences of the V3 loop from many different HIV-1 isolates shows great variation between isolates. Antibodies raised against the V3 loop are

therefore usually type-specific. However, approximately 60% of isolates to date have the consensus sequence GPGRAF, and more than 80% have a GPGR sequence at the tip of the loop. Recent studies have shown that immunisation with peptides containing the GPGRAF consensus sequence or cross-immunisation with recombinant gp120 from different isolates can induce antibodies which cross react between isolates. The GPGRAF consensus sequence may itself be used in the invention.

Other embodiments of the invention involve the use of short sequences of V3 which are not necessarily conserved between various isolates. Whatever V3-derived or V3-related sequence is used, the resulting fusion proteins, or at least particles assembled from them, will be similar antigenically to natural V3 loop sequences in the sense that they cross-react with one or more common antibodies.

More than one V3-derived sequence can be present in a fusion protein of the invention. This embodiment may enable a single fusion protein to be useful in the protection against more than one HIV isolate: therefore, V3-derived sequences from different HIV isolates can be present on the same molecule.

More generally, it will be appreciated that the invention provides considerable flexibility in the nature of the antigenic, second amino acid sequences and the way in which they (if there are more than one) can be located within the first amino acid sequence. For example, two or more identical second amino acid sequences can be inserted in tandem into the same insertion site, two or more identical second amino acid sequences can be inserted into different insertion sites, two or more different second amino acid sequences can be inserted in different insertion sites (or even a single insertion site), and it will be appreciated that the two or more different amino acid sequences may be derived from different epitopes of the same antigen.

As fusion proteins in accordance with the invention spontaneously assemble into particles, it is possible by means of the invention to prepare multivalent particles.

According to a second aspect of the invention, there is provided a particle comprising a plurality of non-natural proteins as described above. Particles in accordance with the invention may contain a heterologous, or, preferably,

homologous population of proteins. Each protein may have any of the configurations described above.

According to a third aspect of the invention, there is provided nucleic acid (particularly DNA) coding for a fusion protein as described above. It will generally be the case that the nucleic acid will be capable of being expressed without splicing or anti-termination events. There will generally be no frame shifting, but frame shifting is not necessarily always excluded.

Further according to the present invention is provided a vector including nucleic acid as described above.

Expression vectors in accordance with the invention will usually contain a promoter. The nature of the promoter will depend upon the intended host expression cell. For yeast, *PGK* is a preferred promoter, but any other suitable promoter may be used if necessary or desirable. Examples include *GAPD*, *GAL*1-10, *PH05*, *ADH*1, *CYC*1, Ty delta sequence, *PYK* and hybrid promoters made from components from more than one promoter (such as those listed). For insect cells, preferred promoters are the polyhedrin and p10 promoters from *Autographica californica* nuclear polyhedrosis virus (AcNPV). Those skilled in the art will be able to determine other appropriate promoters adapted for expression in these or other cells. Vectors not including promoters may be useful as cloning vectors, rather than expression vectors.

The invention also includes host cells, for examples bacterial cells, such as *E. coli*, which may be used for genetic manipulation, yeast cells such as *Saccharomyces cerevisiae* or *Pichia pastoris* or animal cells.

The augmented immunogenic nature of the particles in accordance with the invention, facilitates the production of antibodies with specific characteristics. The invention thus further provides antibodies raised or directed against particulate antigens of the invention; such antibodies may be polyclonal or monoclonal. For the production of human monoclonal antibodies, hybridoma cells may be prepared by fusing spleen cells from an immunised animal with a tumour cell. Appropriately secreting hybridoma cells may thereafter be selected.

Particulate antigens in accordance with the invention may be used in the preparation of vaccines, for example immunotherapeutic vaccines, which form a further aspect of the invention. The vaccine may comprise a particulate antigen and a physiologically acceptable non-toxic carrier, such as sterile physiological saline or sterile PBS. Sterility will generally be essential for parenterally administrable vaccines. One or more appropriate adjuvants may also be present, but are not always necessary. Examples of suitable adjuvants include muramyl dipeptide compounds such as prototype muramyl dipeptide, aluminium hydroxide and saponin.

Vaccines in accordance with the invention may present more than one antigen. Either a cocktail of different particulate antigens may be used, or a homogeneous population of particulate antigens having more than one epitope could be used, as described above. It may in practice be simpler for a vaccine to contain a mixture of different particulate antigens.

Fusion protein and particulate antigens of this invention are useful as diagnostic reagents. Particulate antigens for diagnostic purposes are particularly advantageous because they can be physically separated by centrifugation or filtration and can be directly dispersed on solid supports such as glass or plastic slides, dip sticks, macro or micro beads, test tubes, wells of microtitre plates and the like. The particulate antigens of this invention may also be dispersed in fibrous or bibulous materials such as absorbent disk (US-A-4,632,901), strips or chromatography columns as the solid support. The particles and fusion proteins readily adhere to solid supports. The particles may after purification be disrupted into fusion proteins and the fusion proteins may be dispersed on surfaces as indicated above. These reagents are useful for a variety of diagnostic tests. For example, a test sample suspected of having antibody to the particulate antigen and fluorescent, enzyme or radio-labelled antibody is competitively reacted with the particulate antigen or fusion protein on a solid support and the amount of labelled antibody which binds to the particulate antigen on the solid support. Particulate antigens of this invention are also useful for agglutination reactions with antibodies. Those skilled in the diagnostic arts will recognise a wide variety of application of particulate antigens and fusion proteins of this invention for diagnostic purposes.

Preferred features for each aspect of the invention are as for the first aspect *mutatis mutandis*.

The following examples illustrate the invention, but are not intended to limit the scope in any way. The examples refer to the accompanying drawings, in which;

Figure 1 shows Pepscan analysis of mouse sera. Each plot shows $OD_{492}$ (abscissa) versus peptide number (ordinate) from 1 at the N-terminus to 187 at the truncated C terminus of p1, showing the reactivity of each peptide to antibodies in the test serum. Each test serum is from the pooled sera of five inbred mice immunised with OGS200 VLPs ( described below) in the indicated adjuvant.

Figure 1a: RIBI; 1b: SAF-1; 1c: Chemivax
1d: normal mouse serum 1e: Alum 1f: unadjuvanted

Figure 2 shows data from pre-absorption experiments used to determine epitope surface accessability in three separate rats. The upper plots show Pepscan activities in sera from rats immunised with OGS200 VLPs in alum. The lower plots show the same sera after preincubation with MA5260 VLPs at 4°C overnight and Pepscan analysis. The loss of reactivity with the peptides is due to sequestration of antibodies by epitopes at the surface of the native VLP. The loss of reactivity is specific for certain epitopes.

Figure 3 shows a summary of surface accessibility of epitopes of p1. The sera used for this summary are from two rabbits immunised with MA5260 VLPs in Freunds and three rats immunised with OGS200 VLPs in alum. The discontinuous bar represents those areas of the p1 protein recognised in the pepscan analysis by antibodies in these sera.

Figure 4 shows the location of the insertion sites A, B, C1 and C2 within the regions A, B and C of p1, as defined by the reactive peptides in the Pepscan analysis. The numbers at the end of each sequence are the p1 amino acid coordinates.

Figure 5 shows plasmid pOGS440.

**Example 1 Identification of Epitopes in p1**

The PEPSCAN™ kit (CRB, Cambridge) which was prepared for Ty comprises 10-mer peptides overlapping by 8 residues corresponding to the entire length of the wild type p1 protein of Ty1. 187 peptides cover the truncated p1 protein. Each well of a microtitre plate was coated with a peptide and the anti-Ty test serum overlaid. Antibody binding to epitope peptides was detected by a secondary antibody conjugate and a colorimetric reaction.

Sera from five species (human, macaque, rabbit, rat and mouse) were obtained following immunisation with a variety of VLPs (OGS200: p1-HIVp24 (disclosed in WO-A-8803562), MA5620: p1 alone (disclosed in WO-A-8803563), OGS561 : p1-IIIB:MN:RF V3 loops and OGS530.
OGS 561 is a derivative of pOGS 40, which is disclosed in copending patent application PCT/GB92/01545. At the 3'end of TyA gene are three consecutive V3 loops, in order HXBII, MN, RF. These comprise the amino acid sequences
SNCTRPNNNTRKRIRIQRGPGRAFVTIGKIGMMRQAHCNISG (SEQ ID 1)
SNCTRPNYNKRKRIHIGPGRAFYTTKNIIGTIRQAHCNISG (SEQ ID 2)
SNCTRPNNNTRKSITKGPGRVIYATGQIIGDIRKAHCNLSGS (SEQ ID 3)
which are linked by Bam H1 sites which encode two redundant amino acids glycine and serine. The corresponding nucleotide sequences could readily be determined by persons skilled in the art.

pOGS 530 (and pOGS 531 discussed below) are derivatives of pOGS 40, which is disclosed in copending patent application PCT/GB92/01545. These have an oligonucleotide insertion in the Bam H1 site which encodes the MN (Example 10) or HXBII V3 loop respectively. The immunisations were carried out in different adjuvants (alum, RIBI DETOX™, CHEMIVAX™, SAF-1 or Freund's complete). Sera were analysed by PEPSCAN™. Figure 1 shows a typical raw data set from pooled groups of five mice immunised with OGS200 VLPs in different adjuvants. A summary of peptides recognised by all the sera tested is collated in Table 1. The number of epitopes is, to some extent, adjuvant dependent. A summary of the mouse data from Figure 1 is shown in Table 2 to illustrate this dependence by comparing no adjuvant, alum, CHEMIVAX, RIBI and SAF-1. The use of any of the four adjuvants elicits antibodies to more epitopes than no adjuvant. SAF-1

causes antibodies to be raised to more epitopes (8) than RIBI (5), CHEMIVAX (4) and alum (4). A similar effect has been seen in rabbits. Sera from rabbits immunised with OGS200 VLPs in alum recognised a total of 8 p1 epitopes, whereas with SAF-1 12 epitopes were recognised (Table 1). Freund's appears to be the most powerful adjuvant. Nineteen p1 epitopes were recognised by sera from rabbits immunised with OGS5620 VLPs in Freund's (Table 1)

The choice of epitopes for engineering is extensive; however three "consensus" epitopes emerge from the data. These are contained within peptides 11-17, 28-33 and a larger region covered by peptides 47-68. These correspond to amino acid residues 21-42, 55-74 and 93-142 of the p1 protein and have been named A, B and C, respectively. They are recognised by the overwhelming majority of sera, irrespective of the immunising VLP and the adjuvanting regime.

**Table 1** shows the reactivity of animal sera to p1 peptides in the Pepscan analysis. Each cell in the table shows the number of responders over background, blanks indicate no response. Of the 16 human clinical trial sera tested, only one had a sufficiently high anti-Ty titre to give reliable reactivities in the Pepscan analysis. Eliminating the remaining 15 non-responders, the maximum possible score in the total column is 33. The three 'consensus' epitopes, A, B and C correspond to the peptides 11-17, 28-33 and 47-68, respectively.

**Table 2** shows the serum reactivities of mice immunised with OGS200 VLPs in a variety of adjuvants. All immunisations were intramuscular. The shaded rows correspond to the three "consensus" epitopes, A, B and C at peptides 11-17, 28-33 and 47-68.

## Table 1

| Peptide Number | Human | Macaque | Rat | Rat | Mouse | Rabbit | Rabbit | Rabbit | Rat | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| | OGS200 | OGS200 | OGS200 | OGS561 | OGS200 | OGS200 | OGS200 | MA5620 | MA5620 | |
| | Alum | Alum | Alum | Alum | Various | Alum | SAF-1 | Freunds | Alum | |
| 1-6 | | 4/4 | | 2/5 | 1/5 | 4/5 | 4/5 | 2/2 | | 17 |
| 11-17 | 1/16 | 4/4 | 5/5 | 5/5 | 5/5 | | 5/5 | 2/2 | 1/1 | 28 |
| 21-23 | | | | | 3/5 | 1/5 | | | | 4 |
| 24-27 | | 3/4 | | 2/5 | | 2/5 | 1/5 | 2/2 | | 10 |
| 28-33 | 1/16 | 3/4 | 5/5 | 5/5 | 5/5 | 4/5 | 5/5 | 2/2 | 1/1 | 31 |
| 37-42 | 1/16 | 3/4 | | | 2/5 | | | 1/2 | | 7 |
| 47-68 | 1/16 | 4/4 | 5/5 | 5/5 | 5/5 | 5/5 | 5/5 | 2/2 | 1/1 | 33 |
| 72-77 | | 2/4 | | 2/5 | | 3/5 | 4/5 | 1/2 | | 12 |
| 81-85 | 1/16 | 3/4 | 1/5 | 2/5 | 1/5 | | 1/5 | 1/2 | | 10 |
| 88-92 | 1/16 | | | 3/5 | 2/5 | | | 1/2 | | 7 |
| 101-102 | | 1/4 | | | | | | 2/2 | | 3 |
| 112-118 | | 4/4 | | 1/5 | | | 4/5 | 2/2 | | 11 |
| 128-134 | | 4/4 | 3/5 | | | | 4/5 | 2/2 | 1/1 | 14 |
| 139-140 | 1/16 | 4/4 | 1/5 | 1/5 | | | | 2/2 | | 9 |
| 144-146 | | | 2/5 | 4/5 | | | 1/5 | | | 7 |
| 150-151 | | | 3/5 | | | | | 2/2 | | 5 |
| 157-161 | | 1/4 | | 1/5 | | | 1/5 | 1/2 | 1/1 | 5 |
| 169-171 | | 1/4 | 1/5 | | | | | 2/2 | | 4 |
| 175-182 | | 3/4 | 5/5 | 2/5 | | 1/5 | 4/5 | 2/2 | 1/1 | 18 |
| 185-187 | | | | | | 1/5 | | 1/2 | | 2 |

Table 2

| Peptide Number | ADJUVANT | | | | |
|---|---|---|---|---|---|
| | None | Alum | Ribi | Chemivax | SAF-1 |
| 1-6 | - | - | - | - | + |
| 11-17 | + | + | + | + | + |
| 21-23 | - | - | + | + | + |
| 28-33 | + | + | + | + | + |
| 37-42 | - | + | - | - | + |
| 47-68 | + | + | + | + | + |
| 81-85 | - | - | - | - | + |
| 88-92 | - | - | + | - | + |

**Example 2 Identification of Surface Epitopes of p1**

PEPSCAN™ analysis will identify any well defined linear epitope of p1. Since the analysis is based on recognition of short linear peptides, conformational or non-contiguous epitopic determinants are unlikely to be detected. In addition, PEPSCAN™ data do not distinguish between surface (*ie* accessible to the antibody) or buried epitopes of the native VLP.

Serum preabsorption studies were used to determine which regions of p1, and in particular which of the three epitopes identified above, are surface accessible. Sera from three rats immunised with OGS200 VLPs in alum and from two rabbits immunised with MA5620 VLPs in Freund's were incubated with native purified MA5620 VLPs. These sera were then analysed by PEPSCAN™. Antibodies to surface accessible epitopes bind to the surface of the native VLP and are therefore unavailable to bind to the PEPSCAN™ peptides. Where an epitope is surface accessible, a loss of previously observed reactivity with that epitope indicates that it is a surface feature. The preabsorption experiments were controlled for proteolysis of the native VLP by serum proteases by analysing the particles post-absorption by western blot.

Figure 2 shows the PEPSCAN™ data from three rat sera before and after preabsorption. These data show that epitopes in the N terminal half of p1 are mostly surface accessible whereas those in the C terminal half of the protein are mainly inaccessible. The three major linear epitopes, A, B and C identified above, all showed surface accessibility. A diagrammatic summary of preabsorption experiment data is shown in Figure 3 which illustrates the accessibility of linear epitopes of p1. The gaps are due to regions of the protein not recognised by antibodies in any of the sera tested. This analysis demonstrates that surface accessibility, where information is available, is essentially limited to the N terminal half of p1.

**Example 3 Choice of p1 Epitopes for Engineering**

The three consensus epitopes identified satisfy several criteria for selection as targets for insertion of antigen: they are recognised by sera from all species tested irrespective of the VLP type used as an immunogen and the adjuvanting regime; and all are surface exposed.

Four insertion points within the p1 protein were chosen, one each in A and B, and two in C. These lie between amino acids 30-31, 67-68, 113-114 and 132-133 and are referred to as A, B, C1 and C2 respectively (see Figure 4). Although these four sites were chosen for evaluation, other positions within the defined regions A, B and C may be equally appropriate as insertion sites.

**Example 4 Manipulation of the TYA(d) Gene**

The TyA (d) gene was manipulated to introduce a unique *Nhe* I restriction site at insetion points A, B, C1 or C2 to allow insertion of foreign DNA sequences. Four versions were thus constructed, one for each of the four insertion points selected. The vectors containing this modification were prepared as follows. A *BglII/BamH1* restriction fragment containing the coding sequence of the TyA(d) gene was excised from pOGS 226 and inserted into the vector pSP46 also digested with *BglII/BamH1*, to give pOGS460 (pSP46 is a derivative of pSP64 in which the *HindIII* site in the polylinker has been converted to a *BglII* site).
pOGS 460 was then digested with *NheI* (restriction site present within pSP46) and *PstI* (restriction site present within TyA gene) to release a 1117bp fragment. This was then inserted into M13 mp18 digested with XbaI and PstI. Using site directed mutagenesis, an *Nhe* I restriction site was then introduced at the insertion points A, B, C1 or C2 (ie between TyA nucleotides 90-91, 201-202, 339-340 and 396-397 respectively)

The *Nhe*I site was used for insertion of double stranded (ds) oligonucleotides encoding each of three size variants of the MN isolate V3 loop.

The mutagenised TyA(d) sequences were removed from M13 as *BglII/SplI* fragments and ligated into the vector backbone of *BglII/SplI* digested pOGS440. The *SplI* restriction site in the TyA (d) gene is 5' to the *Pst I* site.

These manipulations yielded the following plasmid constructions:

| | | |
|---|---|---|
| pOGS810 | is the pOGS440 equivalent with the *Nhe*I site at position | A |
| pOGS811 | " | B |
| pOGS812 | " | $C_1$ |
| pOGS813 | " | $C_2$ |

pOGS440 was constructed as follows. pKV560 is described by Chambers *et al.*, (1989) Mol. Cell. Bio. 9 5516-5524. pKV572 is identical to pKV560 with the exception that the interferon sequences are removed leaving a *BglII* cloning site. pKV572 contains the minimal assay promoter with a 5' *BamH1* cloning site for upstream activating sequences, and is the starting point for pJC87.

A 1kb EcoR1-Xho1 fragment from pUG4IS containing the GAL-10 promoter sequence was purified. This was further digested with *Dde1* and a 510 base pair fragment isolated. The 5' protruding ends of this fragment were filled-in with the Klenow fragment of DNA polymerase and BglII oligonucleotide linkers added.

The fragment was then digested with *Sau3A* and a 360 base pair fragment purified. This fragment was ligated into BamH1 digested, phosphate treated pKV572. The ligated products were transformed into HW87 and the resultant plasmids screened for the orientation of the insert by DNA sequencing. A clone which had the 360 base pair GAL1-10 *Dde1-Sau3A* fragment in the GAL1 orientation was selected and called pJC78.

pOGS440 is shown in Figure 5; it was constructed by inserting the BglII/SalI fragment from pOGS226 (a derivative of pMA5620 described in WO-A-8803563 which has an additional BglII site inserted adjacent to the N-terminus of p1) into BglII/SalI at pJC78.

**Example 5 Particle Formation by Insertion Site Mutants of p1**

Insertion of an *Nhe*I restriction site into the TyA(d) gene as described in Example 4 resulted in the introduction of two additional amino acids (Alanine and Serine) into the p1 protein. It was necessary to confirm that this change did not interfere with particle formation for any of the chosen insertion sites (A, B, C1 or C2).

Plasmids pOGS810, pOGS811, pOGS812 and pOGS813 were transformed into *S. cerevisiae* strain MC2, although any available strain could be used. The transformed cells were cultured, harvested and the VLPs isolated by fractionation on sucrose gradients as follows.

Yeast cells were grown selectively at 30°C to a density of $8x10^6$ cells/ml. The cells were then collected by low speed centrifugation, washed once in ice-cold water and resuspended in TEN buffer (10mM Tris, pH 7.4, 2mM EDTA, 140mM NaCl) at 1ml per 1 litre of cells. The cells were disrupted by vortexing with glass beads (40-mesh, BDH), at 4°C until >70% were broken. The beads were pelleted by low speed centrifugation (2,000g), then the supernatant was collected, and the debris removed by centrifugation at 13, 000g for 20 minutes.

The clarified supernatant was transferrred to a SW28 tube and underlayered with 3ml of 60% w/v sucrose solution in TEN. Tubes were then centrifuged at 28K rpm for 90minutes to band the VLPs at the sucrose interface.

VLPs were recovered and dialysed against TEN to remove the sucrose, then purified further by banding on a pre-formed linear (10-60%) sucrose gradient (SW41 tubes centrifuged at 25krpm for 6 hours). The VLPs were recovered, dialysed and concentrated.

All four constructions expressed particulate p1 protein at levels comparable to the positive control for the experiment, pOGS440, demonstrating that addition of the two residues at the insertion points does not adversely affect particle formation.

## Example 6 Insertion of Antigen (GPGRAF)$_3$

Complementary pairs of DNA oligonucleotides were synthesised encoding the central six residues (GPGRAF) of the gp120 V3 loop from the MN isolate. These are

(5'CTAGTGGTCCAGGTAGAGCTTTCT3')$_3$ (SEQ ID 4)

The ends of the annealed double stranded oligonucleotide are compatible with *Nhe*I cut ends for ligation into the unique *Nhe*I sites within pOGS810-813. Transformants were initially screened for the absence of the *Nhe*I site which is abolished by oligonucleotide insertion before DNA sequencing for verification. Three tandemly repeated copies were inserted at position B in the TYA(d) gene of pOGS811, to generate pOGS814. The coding consequence of this is as follows:

P E N P A S **G P G R A F** S S **G P G R A F** S S **G P G R A F** S S H H A S P (SEQ ID 5)

The residues in bold are the inserted amino acids flanked by the wild type p1 residues. The A S and S S motifs are encoded by the *Nhe*I cohesive ends of the oligonucleotide.

Construction of the (GPGRAF)$_3$ insert has provided information on the size of insert that can be tolerated at position B. Since this 26 residue insertion allows particle formation, insertion of the 20 and 40 residue V3 loop size variants should be tolerated at that position. It also supports the notion that the original Ty epitope is in the form of a surface loop which can be extended without interfering with the normal fold of the p1 monomer.

**Example 7 Insertion of Antigen: GPGRAF**

Oligonucleotides encoding for GPGRAF with *Nhe*I compatible ends were synthesised as described above.

Once annealed they were ligated into each of the four insertion sites. Once the oligonucleotide was inserted, the *Nhe*I site was abolished. Resulting transformants were therefore screened for loss of the *Nhe*I site. The orientation of the insert was verified by DNA sequencing. The resulting constructions are numbered as follows:

| | | |
|---|---|---|
| pOGS815: | pOGS810 with GPGRAF at position | A |
| pOGS816: | pOGS811 | B |
| pOGS817: | pOGS812 | C1 |
| pOGS818: | pOGS813 | C2 |

The total inserted sequence is as follows:

A S G P G R A F S S (SEQ ID 6)

The AS and SS residues flanking the N and C termini of the inserted antigen respectively are encoded by the altered *Nhe*I sites at each end of the inserted oligonucleotide.

*S. cerevisiae* strain MC2 yeast cells were transformed with each plasmid.

## Example 8 Insertion of Antigen: Half V3 loop

Complementary pairs of DNA oligonucleotides

5'CTAGTAAAAGAA AGAGAATTCATATTGGTCCAGGTAG AGCTTTCTATAC TACCAAAAACATTATCG3' (SEQ ID 7)

were synthesised that encode the following sequence:

A S **K R K R I H I G P G R A F Y T T K N I I** A S (SEQ ID 8)

The flanking AS residues are those encoded by the *Nhe*I compatible oligonucleotide ends. The annealed oligonucleotide possessed an *Eco*RI restriction site. Once ligated into the vector the 5' *Nhe*I site was abolished while the 3' *Nhe*I site was recreated. The remaining 3' *Nhe*I site enables further antigens to be added if desired. Transformants were screened by *Eco*RI restriction digestion and the orientation of insertion was determined by DNA sequencing. The resulting constructions are numbered as follows:

| | | |
|---|---|---|
| pOGS819: | pOGS810 with the half V3 loop at position | A |
| pOGS820: | pOGS811 | B |
| pOGS821: | pOGS812 | C1 |
| pOGS822: | pOGS813 | C2 |

**Example 9 Insertion of Antigen: Whole V3 loop**

Two pairs of complementary pairs of DNA oligonucleotides

5'CTAGT ATTAATTGCACCC GTCCTAACTACA ATAAAAGAAAGAG AATTCA TATTGGTCCAGGT3' (SEQ ID 9) and

5'AGA GCTTTCTA TACTACCAA AAACATTA TCGGTACT ATTAGACAA GCTC ACTGTAATATCG3' (SEQ ID 10)

were synthesised that together encode the whole V3 loop sequence as follows:

A S I N $C^S$ T R P N Y N K R K R I H I G P G R A F Y T T K N I I G T I R Q A H $C^S$ N I A S. (SEQ ID 11)

The flanking AS residues were encoded by the *Nhe*I compatible ends and $C^S$ signifies the cysteine residues thought to close the loop at its base by a disulphide bond. The whole insert was constructed in two parts which were ligated together before ligation into the appropriate vectors. As with the half loop oligonucleotides, the 5' *Nhe*I site is abolished on insertion and the 3' *Nhe*I site is recreated. The inserted sequence also carries an *Eco*RI restriction site to aid screening. The resulting transformants were screened for the presence and orientation of the DNA fragment by restriction enzyme digestion. The three ligation junctions, at each end and in the middle of the insert, were verified by DNA sequencing. The constructions were numbered as follows:

| | | |
|---|---|---|
| pOGS823: | pOGS810 with whole V3 loop at position | A |
| pOGS824: | pOGS811 | B |
| pOGS825: | pOGS812 | C1 |
| pOGS826: | pOGS813 | C2 |

**Example 10 Characterisation of pOGS814 VLPs: $(GPGRAF)_3$ at position B**

Purified pOGS814 DNA was transformed into *S. cerevisiae* strain MC2, although any available strain could be used. Cells were harvested, hand bead-beaten and the cell homogenate clarified by centrifugation at 9K for 20 minutes. 1.5ml of this material was then applied to sucrose gradients (15 to 45% with a 60% cushion) and centrifuged at 40 Krpm for 1.5 hours. The gradients were fractionated and examined by SDS-PAGE. The OGS814 protein sedimented with the characteristics of a VLP in a well defined zone half way down the gradient, well resolved from monomeric protein solutes.

**Example 11 Immunoreactivity of OGS814 VLPs**

Fractions from the gradients described in Example 10 were analysed by western blotting with three antibodies: an anti-Ty polyclonal, DuPont gp120 MAb 9305 which reacts with the V3 loop tip sequence -RIQRGPGRAFVTIGK-, and Dupont gp120 monoclonal 9284, which reacts with the left-hand side of the V3 loop-NNNTRKSIRIQR-. As expected, the OGS814 VLPs reacted with the Ty polyclonal and 9305 MAb, but not with 9284 MAb. The controls were MA5620 VLPs and OGS531 VLPs (whole V3 loop from isolate HXB2 at the C terminus). All the controls had the predicted reactivities. The western blot data are summarised in Table 3.

| VLP | Added Antigen | Antigen position | Antibody | | |
|---|---|---|---|---|---|
| | | | anti-Ty | 9305 | 9284 |
| MA5620 | - | - | + | - | - |
| OGS814 | $(GPGRAF)_3$ | B | + | + | - |
| OGS531 | V3 loop | C terminus | + | + | + |

**Table 3.** Western blot immunoreactivity data from MA5620, OGS814, and OGS531 VLPs with anti-Ty, 9305 and 9284 antibodies.

**Example 12 Surface Exposure of the Antigen in OGS814 VLPs**

The p1 epitope at position B was shown to be surface-exposed in native whole VLPs by its ability to bind its cognate antibody which could then be removed from solution by cosedimentation with the VLP during centrifugation. A similar approach was used to demonstrate that the GPGRAF component of OGS814 at position B is also surface exposed. In this case the cognate antibody was the MAb 9305, shown to recognise OGS814 VLPs. The experiment involved incubation of the VLPs with the MAb, pelleting the VLPs by centrifugation and measuring the amount of unbound MAb left in the supernatant using a V3 peptide ELISA.

In an ELISA for detecting MAbs binding to the 40 amino acid HXB2 gp120 V3 loop peptide, VLPs at 100 and 500 µg/ml or peptide at 200 µg/ml were incubated with 9305 or 9284 MAb at a dilution of 1/100 from the stock. Controls for binding in solution were MA5620 VLPs (negative) and the V3 peptide (positive). The mixtures were centrifuged at 75 Krpm for 15 minutes and the supernatants assayed for residual MAb by ELISA. In summary:

1) the MAbs alone were not removed from solution by centrifugation
2) the negative control MA5620 VLPs bound no antibody, which remained in the supernatant
3) the positive control peptide removed all antibody reactivity from the supernatant, *ie* no unbound antibody remained
4) OGS814 VLPs bound 9305, but not 9284 antibodies, indicating that the GPGRAF motif in these VLPs is surface accessible

**Example 13 Immunogenicity of OGS814 VLPs**

Rats were immunised with purified OGS814 VLPs. Rats were primed at week 0, boosted at weeks 6 and 12, and final bleeds were taken at week 14. Intermediate test bleeds were taken at weeks 6, 8 and 12. Two doses of 50 and 250 µg per immunisation, in the presence and absence of adjuvant are given to four groups of five rats as follows:

| | |
|---|---|
| Group 1 | 50 µg - adjuvant/animal |
| Group 2 | 50 µg + adjuvant/animal |
| Group 3 | 250 µg - adjuvant/animal |
| Group 4 | 250 µg + adjuvant/animal |

**Example 14 Immunogenicity of OGS822 VLPs**

OGS822 VLPs were chosen to examine the improved immunogenicity resulting from the insertion of the half V3 loop within the Ty p1 protein. Rats were immunised intramuscularly with 250μg purified OGS822 VLPs in aluminium hydroxide adjuvant. Rats were primed at week 0, boosted at weeks 6 and 12, and final bleeds were taken at week 14. Sera were tested for anti-V3 antibody responses both by ELISA and neutralisation assays, the results of which are shown in Table 4.

| | 4 weeks post-prime | | 2 weeks post boost | |
|---|---|---|---|---|
| ANIMAL | ELISA U/ml | neutralising | ELISA U/ml | neutralising |
| 1 | 21.4 | 320-640 | 28.4 | 1280 |
| 2 | 23.1 | 640 | 44.4 | 1280-2560 |
| 3 | 13.1 | 256 | 11.7 | 320-640 |
| 4 | 8.0 | 256 | 8.2 | 640 |
| 5 | - | 16 | 4.2 | 256 |

**Table 4** Serum antibody and neutralising antibody titres of rats immunised with OGS 822 VLPs. The ELISA data (shown as units/ml) are arbitrary values based on a standard curve produced with a rat anti-MN peptide antiserum. Neutralising antibody titres are expressed as the dilution of serum that resulted in 90% inhibition of syncytia formation in a standard assay.

In the same assays, a pool of antisera from rats immunised with OGS259 VLPs (1/2 V3 loop at the C-terminus) generated an ELISA value of 2.27 U/ml and a neutralisation titre of 1:8. Insertion of antigen at an internal site (in this case C2) thus resulted in a dramatic improvement in immunogenicity.

**Example 15 Insertion of Antigen: Influenza nucleoprotein CTL epitope**

Complementary pairs of DNA oligonucleotides were synthesised that encode the following sequence:

ASRS **TYQRTRALV** GSAS (SEQ ID 12)

This contains an influenza nucleoprotein CTL epitope (shown in bold). The flanking ASRS and GSAS amino acids are encoded by restriction enzyme sites. This sequence was inserted into the p1 protein at each of the four sites A, B, $C_1$ and $C_2$.

## CLAIMS

1 A non-natural particle-forming protein comprising a self-assembling particle-forming first amino acid sequence substantially homologous with a yeast retrotransposon Ty p1 protein and a second amino acid sequence, wherein the second sequence is antigenic and is incorporated within an epitope of the first amino acid sequence, which epitope, on particles formed from the first amino-acid sequence alone, is surface-exposed.

2. A particle-forming protein as claimed in claim 1 wherein the second amino acid sequence is inserted into the surface-exposed epitope.

3. A particle-forming protein as claimed in claim 1 wherein the second amino acid sequence is substituted in place of native amino acids normally present in the surface-exposed epitope.

4. A particle-forming protein as claimed in any of claims 1 to 3 wherein native amino acids normally present in the surface epitope are deleted.

5. A particle-forming protein as claimed in any of claims 1 to 4 wherein the surface-exposed epitope is present in the N-terminal half of the first amino acid sequence.

6. A particle-forming protein as claimed in claim 5 wherein the first amino acid sequence p1 protein of the retrotransposon Ty has been truncated at the C-terminal.

7. A particle-forming protein as claimed in claim 6 wherein the surface-exposed epitope is located between amino acids 21-42.

8. A particle-forming protein as claimed in claim 6 wherein the surface-exposed epitope is located between amino acids 55-74.

9. A particle-forming protein as claimed in claim 6 wherein the surface-exposed epitope is located between amino acids 93-142.

10. A particle-forming protein as claimed in claim 7 wherein the second amino acid sequence is inserted between amino acids 30-31,

11. A particle-forming protein as claimed in claim 8 wherein the second amino acid sequence is inserted between amino acids 67-68

12. A particle-forming protein as claimed in claim 9 wherein the second amino acid sequence is inserted between amino acids 113-114

13. A particle-forming protein as claimed in claim 9 wherein the second amino acid sequnece is inserted between amino acids 132-133

14. A particle-forming protein as claimed in any one of claims 1 to 13 wherein the antigenic second amino acid sequence or sequences correspond to a sequence derived from or associated with an aetiological agent or a tumour.

15. A particle-forming protein as claimed in claim 14 wherein the aetiological agent is a microorganism such as a virus, bacterium, fungus or parasite.

16. A particle-forming protein as claimed in claim 15 wherein the virus is a retrovirus, for example HIV-1, HIV-2, HTLV-I, HTLV-II, HTLV-III, SIV, BIV, ELAV, CIAV, murine leukaemia virus, Moloney murine leukaemia virus, and feline leukaemia virus; an orthomyxovirus, for example influenza A or B; a paramyxovirus, for example parainfluenza virus, mumps, measles, RSV and Sendai virus; a papovavirus, for example HPV; an arenavirus, for example LCMV of humans or mice; a hepadnavirus, for example Hepatitis B virus; or a herpes virus, for example HSV, VZV, CMV, or EBV.

17. A particle-forming protein as claimed in claim 14 wherein the tumour-associated or derived antigen is a proteinaceous human tumour antigen, for example a melanoma-associated antigen, or an epithelial-tumour associated antigen, for example from breast or colon carcinoma.

18. A particle-forming protein as claimed in claim 14 wherein the antigenic sequence is derived from a bacterium, for example of genus *Neisseria*, *Bordetella*, *Listeria*, *Mycobacteria* or *Leishmania*, from a parasite, for example from the genus *Plasmodium*, or from a fungus, for example from the genus *Candida*, *Aspergillus*, *Cryptococcus*, *Histoplasma* or *Blastomyces*.

19. A particle-forming protein as claimed in any of claims 1 to 18 wherein the antigenic sequence is between 6 and 60 amino acids in length.

20. A particle-forming protein as claimed in claim 14 wherein the antigenic sequence is an epitope from:

1) HIV (particularly HIV-1) gp120,
2) HIV (particularly HIV-1) p24,
3) Influenza virus nucleoprotein and haemagglutinin,
4) LCMV nucleoprotein,
5) HPV L1, L2, E4, E6 and E7 proteins,
6) p97 melanoma associated antigen,
7) GA 733-2 epithelial tumour-associated antigen,
8) MUC-1 epithelial tumour-associated antigen,
9) Mycobacterium p6,
10) Malaria CSP or RESA antigens,
11) VZV gpI, gpII or gpIII

21 A particle-forming protein as claimed in claim 20 wherein the epitope is the V3 loop or GPGR loop of the envelope glycoprotein gp120 of a lentivirus.

22. A particle-forming protein as claimed in any of the preceding claims wherein two or more of the surface exposed epitopes have an antigenic amino acid sequence incorporated therein.

23. A particle-forming protein as claimed in claim 22 wherein the antigenic amino acid sequence incorporated within one of the surface-exposed epitopes is different from the antigenic amino acid sequence incorporated within another of the surface-exposed epitopes.

24. A particle-forming protein as claimed in any of the preceding claims wherein more than one antigenic amino acid sequence is incorporated within any single surface exposed epitope.

25. A particle-forming protein as claimed in claim 24 wherein the antigenic amino acid sequences incorporated in any single surface exposed epitope are not all identical.

26. A protein as claimed in claim 24 or claim 17 wherein the antigenic amino acid sequences are incorporated in tandem within the surface exposed epitope.

27. A protein as claimed in claims 23, 25 or 26 where the different second amino acid sequences are derived from different epitopes of the same antigen.

28. A particle comprising a plurality of homologous proteins as claimed in any of claims 1 to 27.

29. A particle comprising a plurality of heterologous proteins as claimed in any of claims 1 to 27 .

30. Nucleic acid coding for a fusion protein as claimed in any one of claims 1 to 27.

31. A vector including nucleic acid as as claimed in claim 30.

32. A host cell carrying a vector as claimed in claim 31.

33. A host cell as claimed in claim 32 where the host cell is *E. coli*

34. A host cell as claimed in claim 32 where the host cell is a yeast cell for example *Saccharomyces cerevisiae* or *Pichia pastoris*

35. Host cells as claimed in claim 32 where the host cell is an insect cell for example *Spodoptera frugiperda*.

36. Antibodies raised or directed against particulate antigens as claimed in any of claims 1 to 29.

37. The use of hybrid proteins and/or particles as claimed in any one of claims 1 to 29 in the preparation of an immunotherapeutic or prophylactic vaccine.

38. The use of particulate antigens as claimed in claims 1 to 29 as a diagnostic agent.

39. A pharmaceutical or veterinary composition comprising a protein as claimed in any one of claims 1 to 29 together with a pharmaceutically and/or veterinarily acceptable carrier.

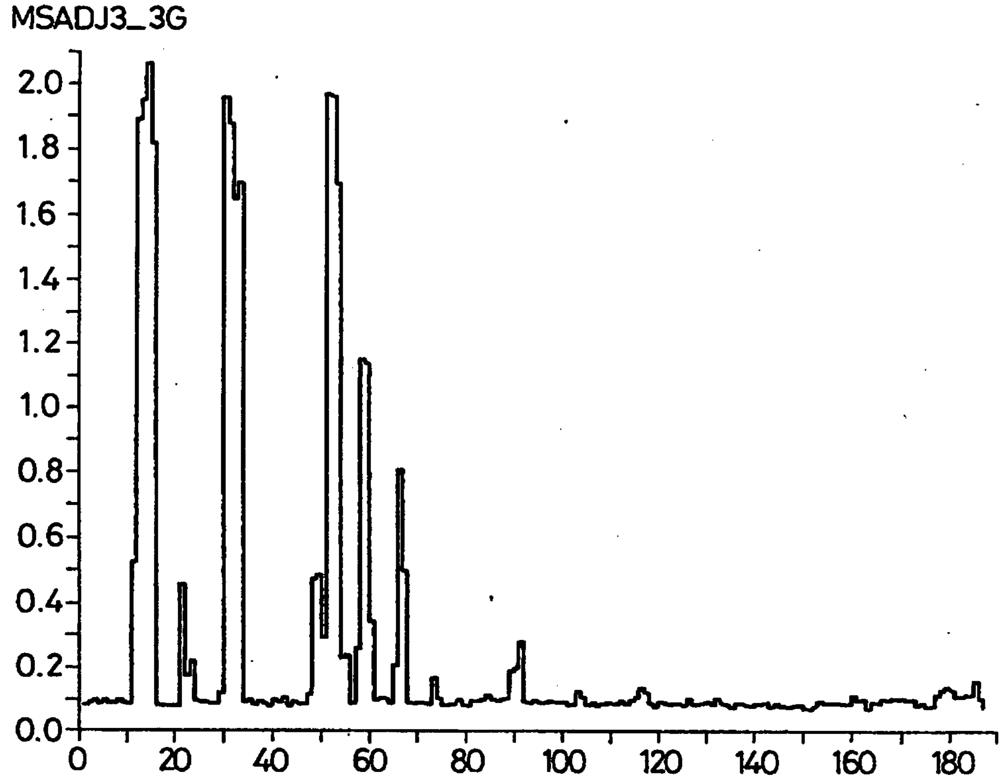
MSADJ3_3G
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
160
180

Fig.1a

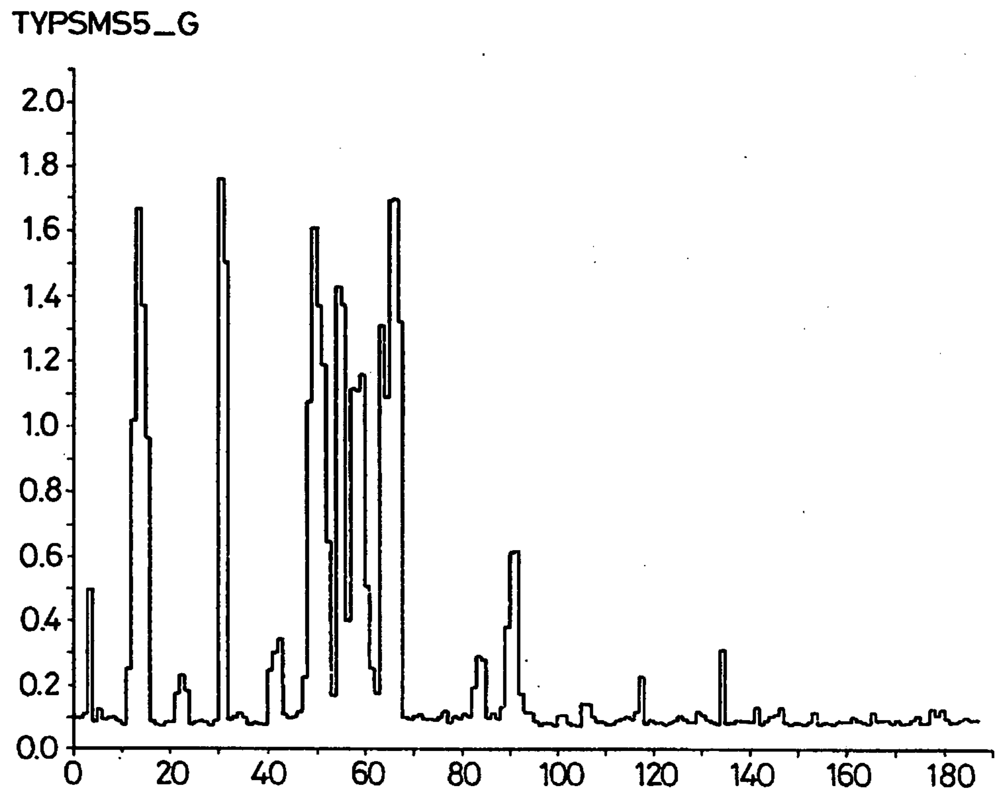
TYPSMS5_G
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
160
180

Fig.1b

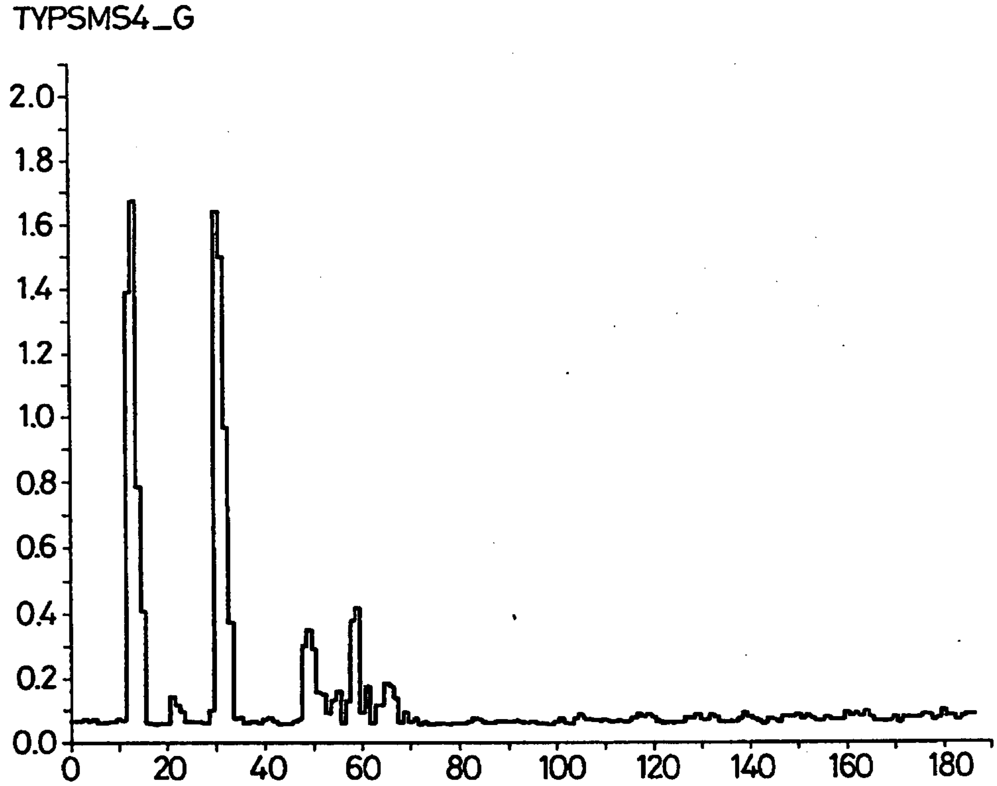
TYPSMS4_G
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
160
180

Fig.1c

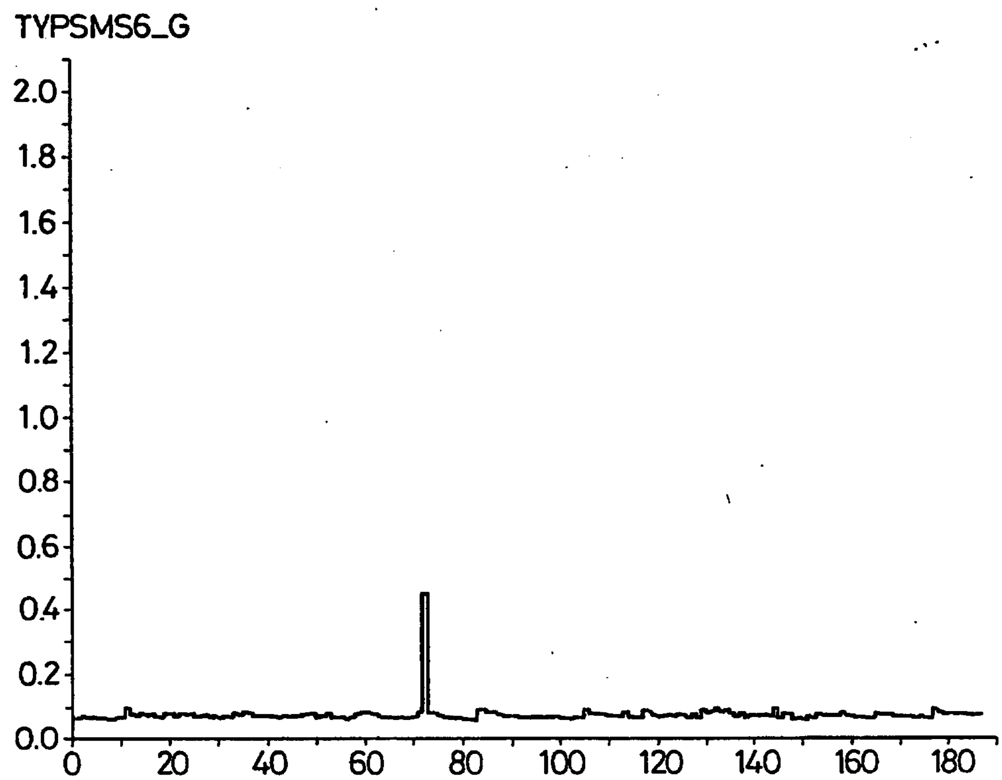
TYPSMS6_G
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
160
180

Fig.1d

3/8

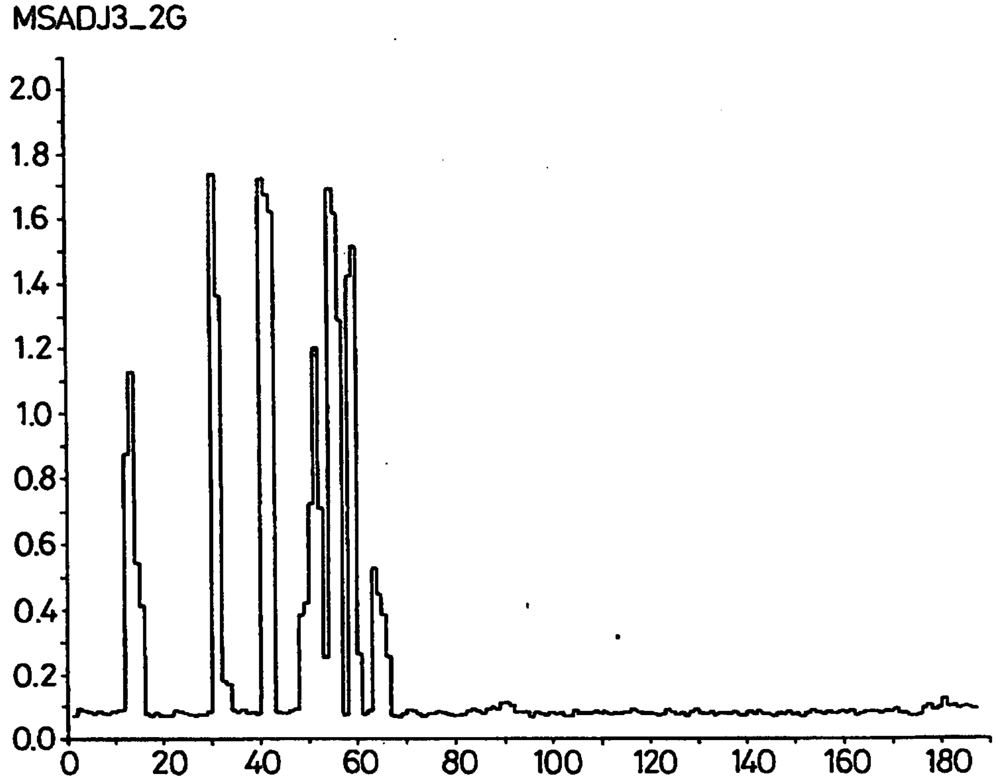


Fig.1e

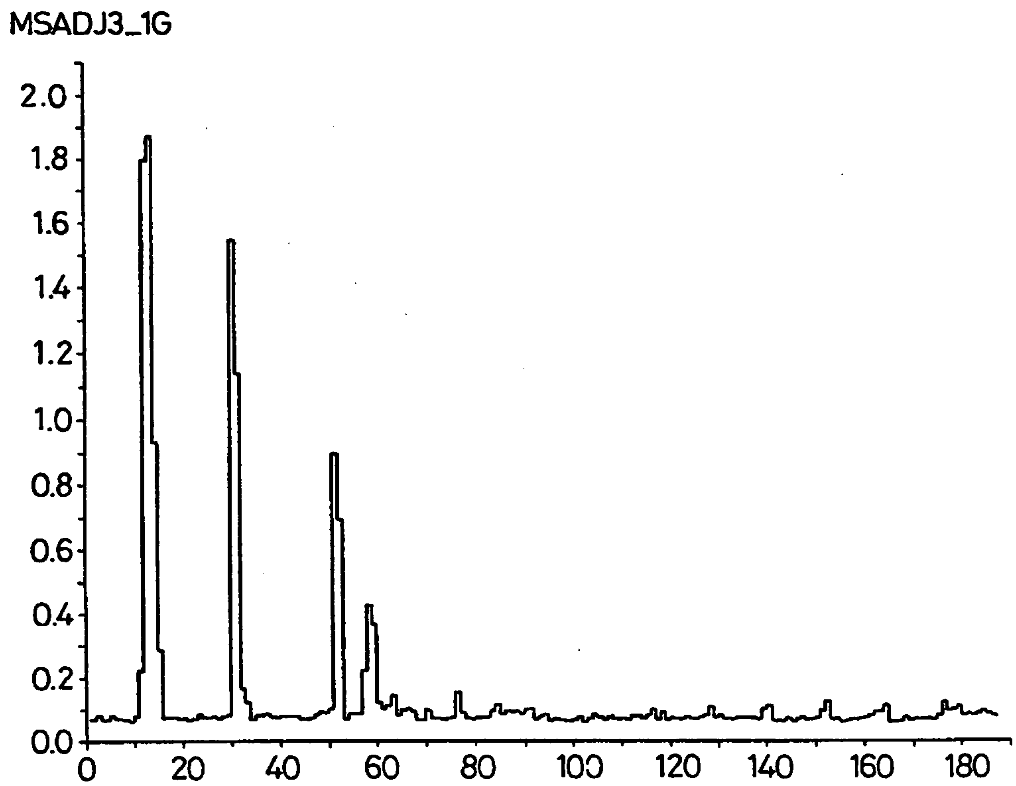


Fig.1f

SUBSTITUTE SHEET

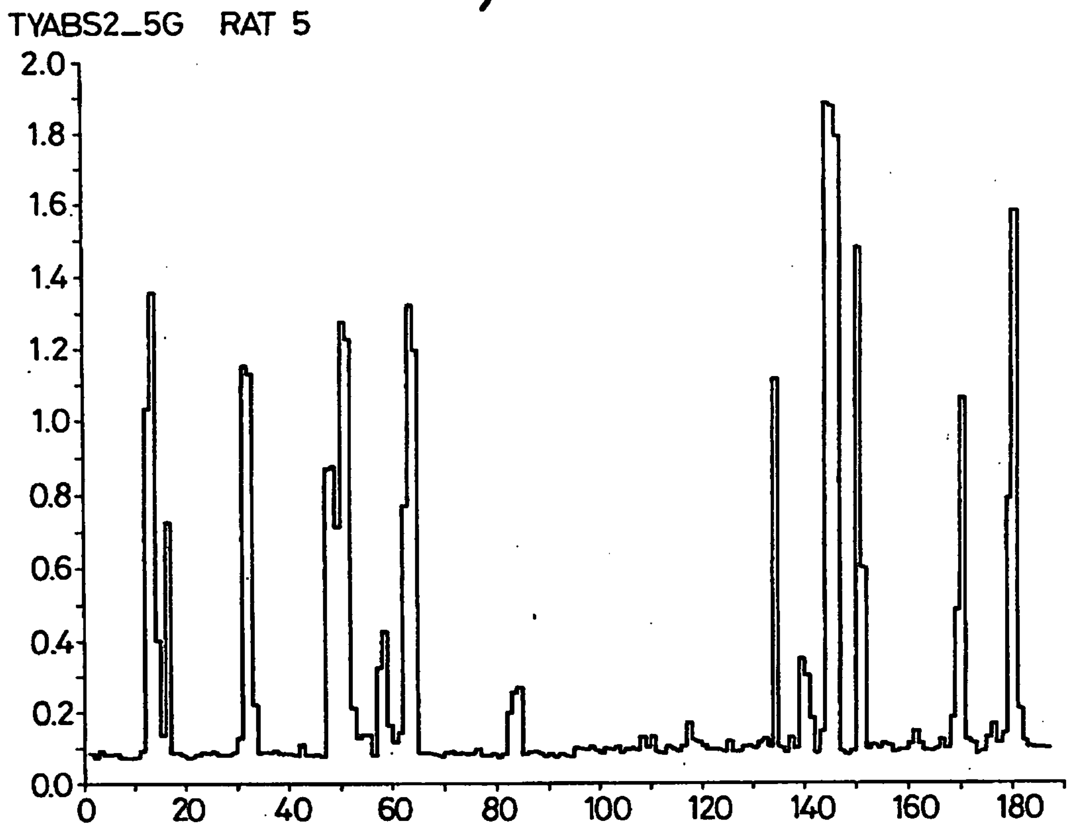


Fig. 2a (I)

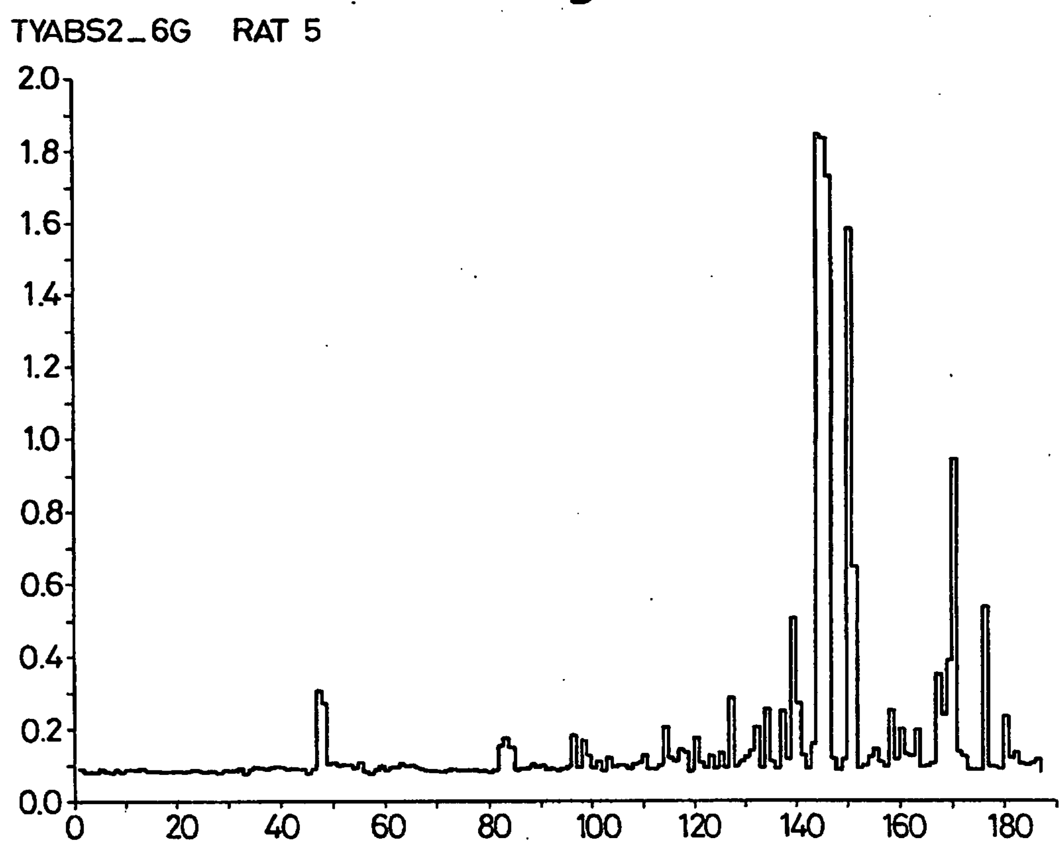


Fig. 2a (II)

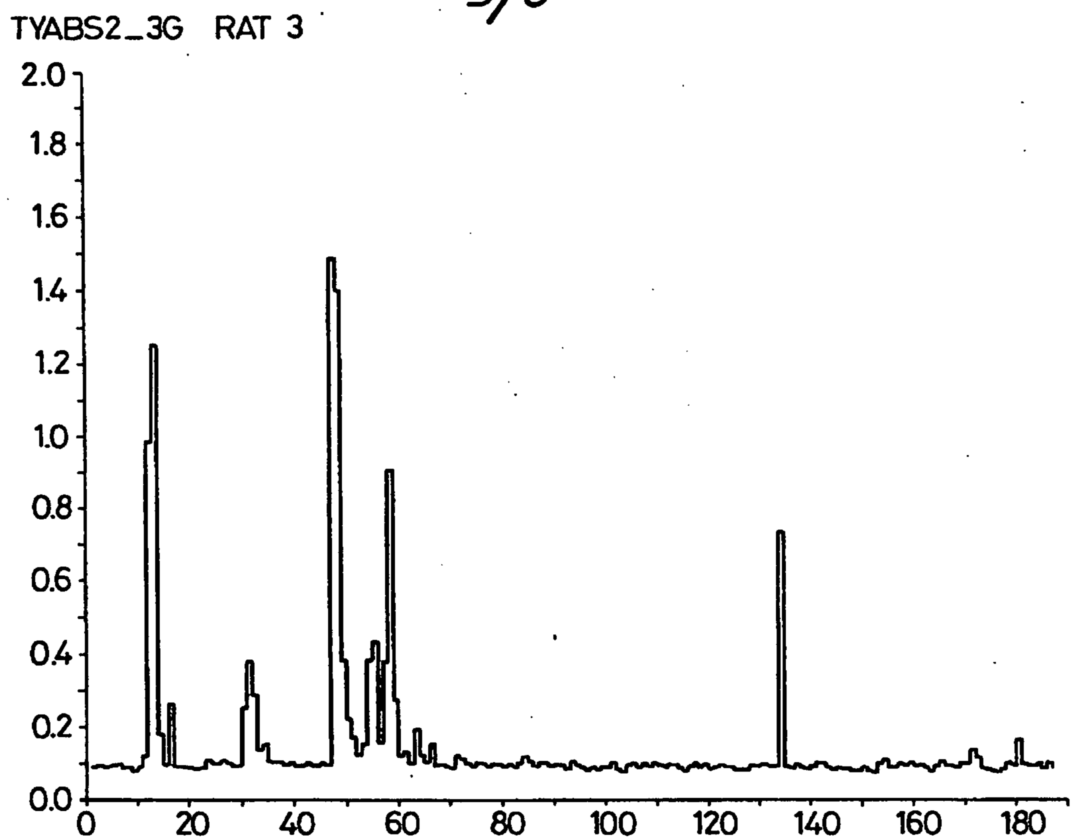


Fig.2b(I)

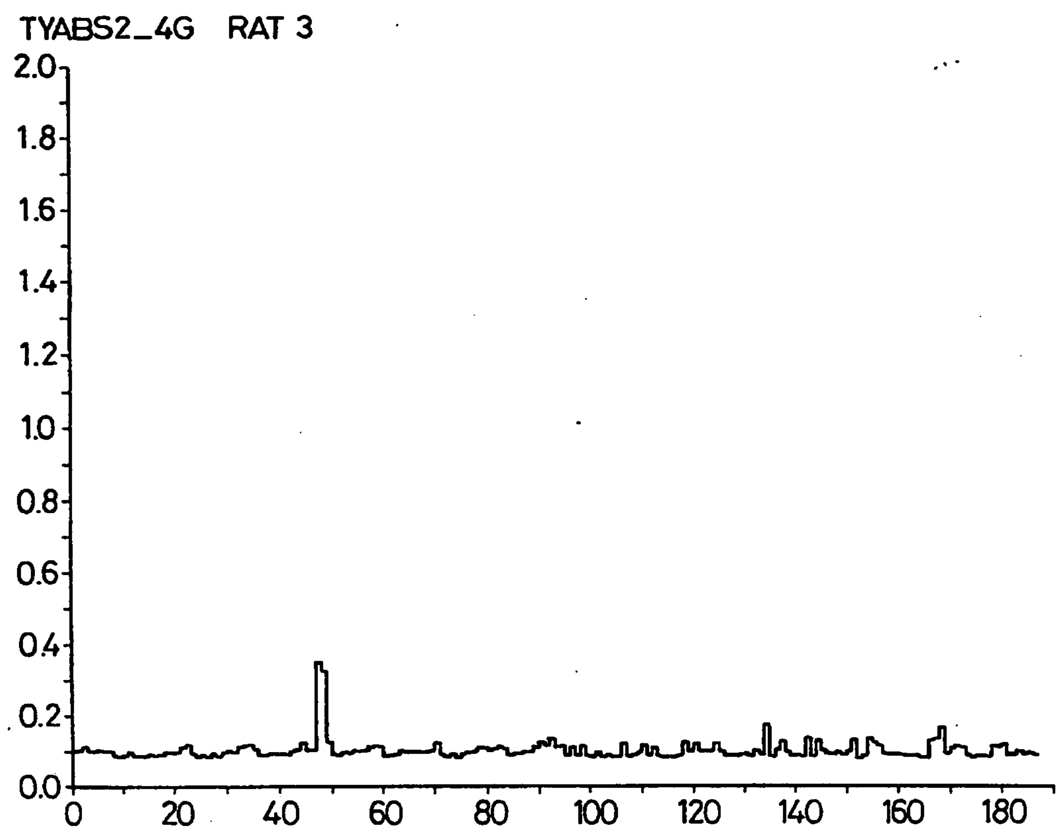


Fig.2b(II)

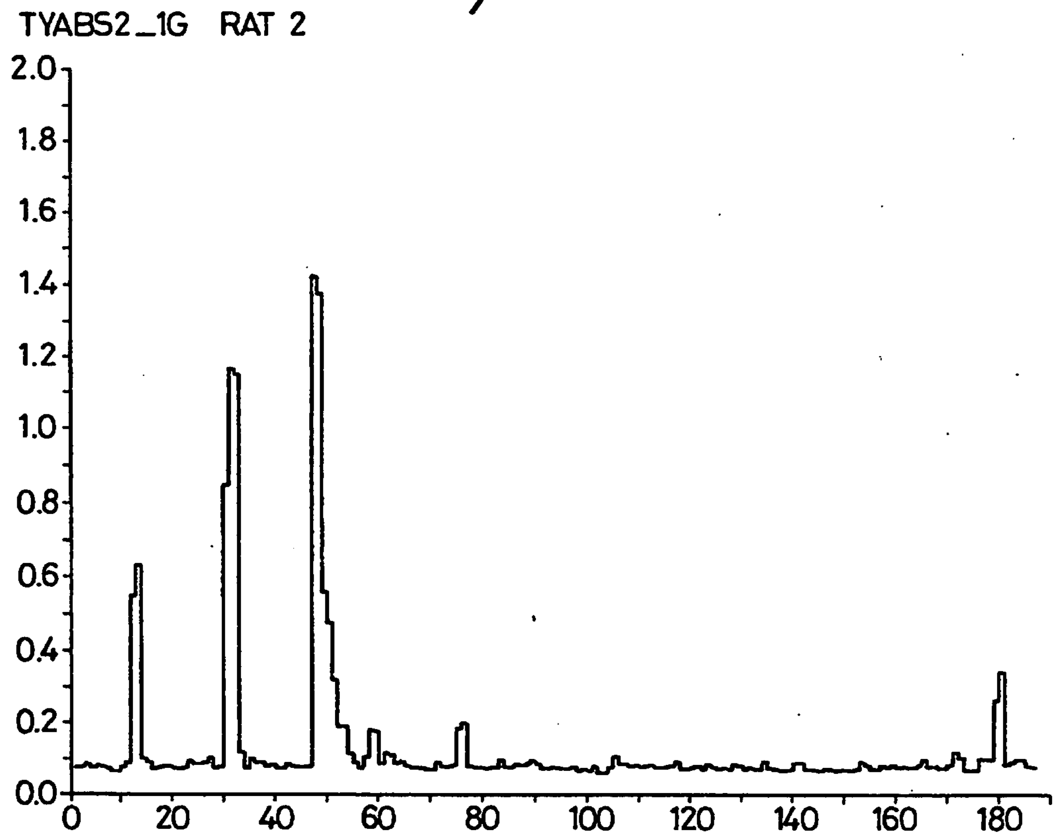


Fig. 2c (I)

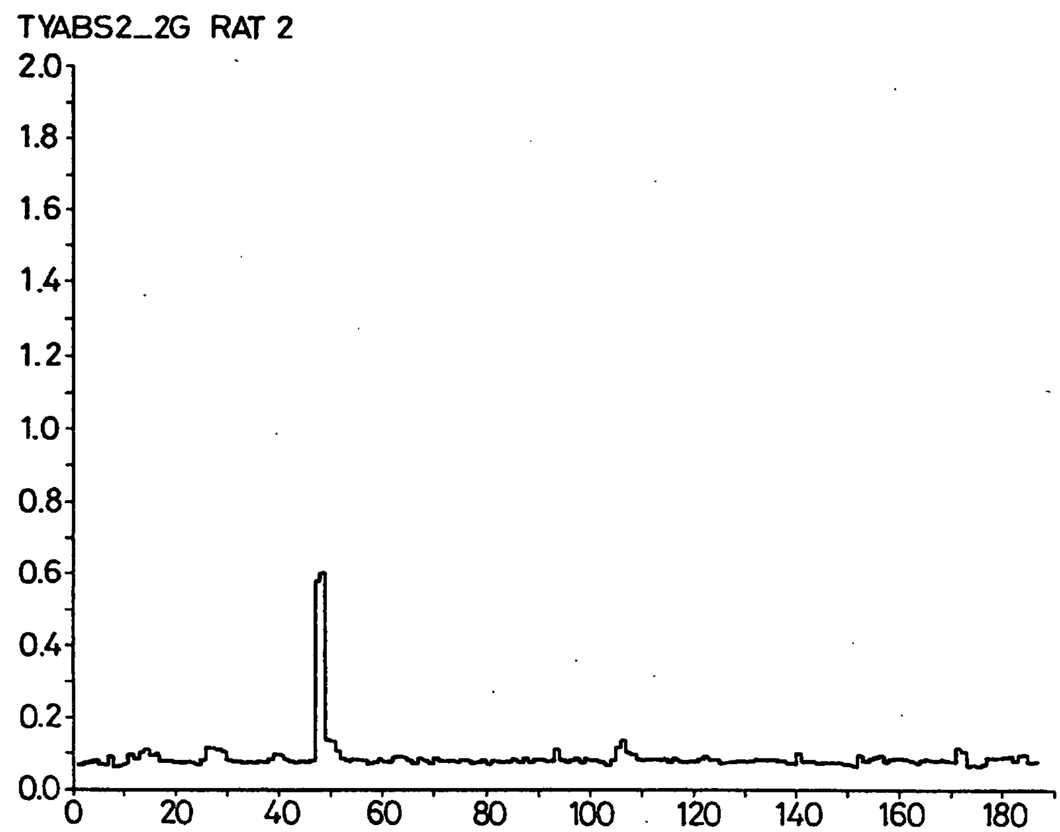


Fig. 2c (II)

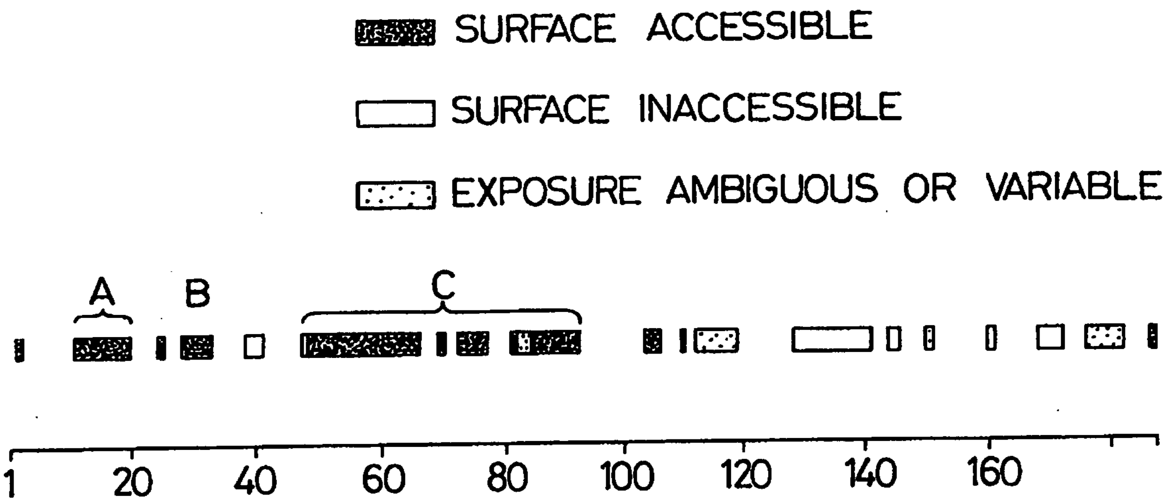


Fig.3

21 42
A VTSKEVHTNQDPLDVSASKTEE
A

55 74
B TTTPASSAVPENPHHASPQT
B

C1
93
C QNQANPSGWSFYGHPSMIPYTPYQMS

C2 142
PMYFPPGPQSQFPQYPSSVGTPLR

Fig.4

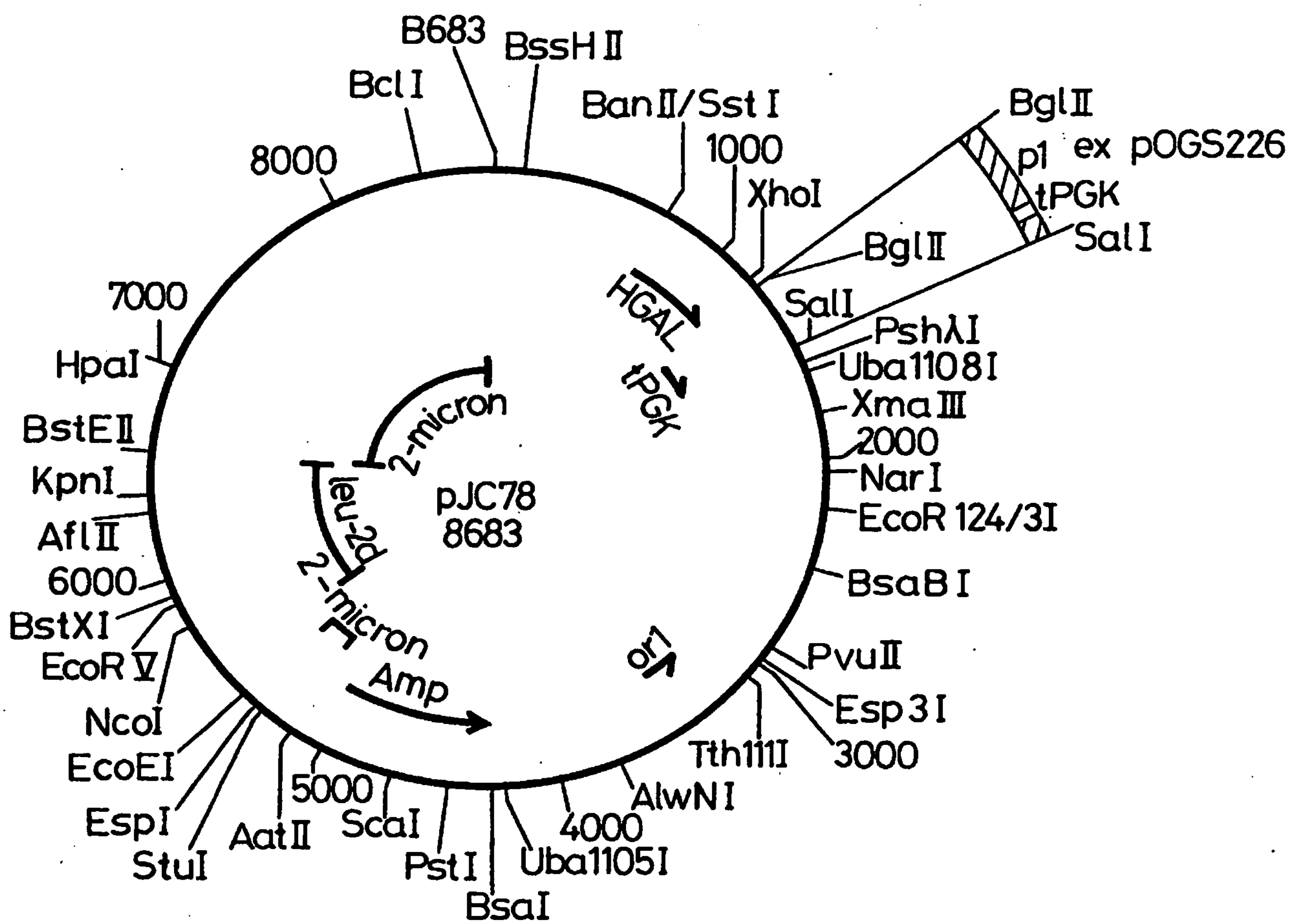
B683
BssH II
Bcl I
Ban II/Sst I
1000
XhoI
Bgl II
p1 ex pOGS226
tPGK
Sal I
Bgl II
8000
7000
HpaI
BstE II
KpnI
Afl II
6000
BstXI
EcoR V
NcoI
EcoEI
EspI
StuI
Aat II
5000
ScaI
PstI
BsaI
Uba1105I
4000
AlwNI
Tth111I
3000
Esp3I
PvuII
BsaBI
EcoR 124/3I
Nar I
2000
Xma III
Uba1108I
PshλI
SalI
HGAL
tPGK
2-micron
leu-2d
2-micron
Amp
ori
pJC78
8683

*Fig.5*

# INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/GB 93/02656

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC 5 C12N15/81 C07K15/00 C12N15/62 C07K13/00 C12N1/21
C12N1/19 C12N5/10 C12P21/08 A61K39/00 G01H33/569

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 5 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | IMMUNOLOGY<br>vol. 77 , 1992<br>pages 315 - 321<br>S.J. HARRIS ET AL.; 'Enhanced proliferative cellular responses to HIV-1 V3 peptide and gp120 following immunization with V3:Ty virus-like particles'<br>*abstract; discussion*<br>--- | 1 |
| A | NATURE<br>vol. 329 , 1987<br>pages 68 - 70<br>S.E. ADAMS ET AL.; 'The expression of hybrid HIV:Ty virus-like particles in yeast'<br>*whole document*<br>--- | 1 |
| | -/-- | |

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 9 May 1994 | 25-05-1994 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax (+31-70) 340-3016 | Yeats, S |

1

Form PCT/ISA/210 (second sheet) (July 1992)

Interna     ıl Application No
PCT/GB 93/02656

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| --- | --- | --- |
| A | SCIENCE<br>vol. 234 , 1986<br>pages 1392 - 1395<br>S.D. PUTNEY ET AL.;<br>'HTLV-III/LAV-neutralizing antibodies to an E. coli-produced fragment of the virus envelope'<br>cited in the application<br>--- | |
| A | THE EMBO JOURNAL<br>vol. 3 , 1984<br>pages 1115 - 1119<br>M.J. DOBSON ET AL.; 'The identification and high level expression of a protein encoded by the yeast Ty element'<br>cited in the application<br>----- | |

1